6.10 Problème de correspondance de Post

Le problème de correspondance de Post (PCP) :

Donnée : Deux suites de mots finies $(u_1, \ldots, u_n)$ et $(v_1, \ldots, v_n)$ de même longueur

Question : existe-t-il un entier k et une suite d'indices $i_1, \ldots, i_k$ tels que

$$u_{i_1} \cdots u_{i_k} = v_{i_1} \cdots v_{i_k}$$

Par exemple : soient

i	1	2	3	4
u_i	a	b	ca	abc
v_i	ab	ca	a	c

Cette instance de PCP a une solution (12314).

Théorème 6.10.1 *PCP est indécidable.*

On commence par montrer que le problème de correspondance de Post *modifié*, dans lequel on fixe $i_1 = 1$ est indécidable.

On réduit le problème de l'arrêt.

Soient M une machine de Turing et $w \in \Sigma^*$. On note q_f le (seul état d'arrêt de M)

Les mots de PCP modifié :

	mots v_i (dans l'ordre)	mots u_i correspondant	
1.	$\lhd$	$\lhd\, q_0\$w\star$	
2.	a	a	pour $a \in \Sigma$
3.	qa	$a'q'$	Si $\delta(q,a) = q', a', \rightarrow$
4.	aqb	$q'ab'$	si $\delta(q,b) = q', b', \leftarrow$
5.	qa	$q'a'$	si $\delta(q,a) = q', a', \downarrow$
6.	$q\star$	$aq'\star$	si $\delta(q,B) = q', a', \rightarrow$
7.	$bq\star$	$q'ba'\star$	si $\delta(q,B) = q', a', \leftarrow$
8.	$q\star$	$q'a\star$	si $\delta(q,B) = q', a', \downarrow$
9.	$\$q_e\star\, \rhd$	$\rhd$	
10.	$\star$	$\star$	
11.	$aq_e\star$	$q_e\star$	
12.	$q_f a$	aq_f	
13.	$q_f\star$	$q_e\star$	

Montrons que PCP a une solution ssi M s'arrête sur w :

Si M s'arrête sur w Soit

$$s_1 = q_0\$w \vdash \cdots \vdash w_k q_f w'_k = s_k$$

le calcul de M sur w (les blancs en fin de ruban n'apparaissent pas dans les configurations). Pour tout $i < k$, s_i, s_{i+1} sont de l'une des formes suivantes :

— $s_i = t_i q a w_i$ et $s_{i+1} = t_i a' q' w_i$ avec $\delta(q, a) = q', a', \rightarrow$
— $s_i = t_i a q b w_i$ et $s_{i+1} = t_i q' a b' w_i$ avec $\delta(q, b) = q', b', \leftarrow$
— $s_i = t_i q a w_i$ et $s_{i+1} = t_i q' a' w_i$ avec $\delta(q, a) = q', a', \downarrow$
— L'un des cas ci-dessus en retirant la partie à droite de q (tête de lecture à droite du ruban)...

Dans chaque cas, on remarque que $s_i\star$ peut s'écrire comme la concaténation $v_{m_1} \cdots v_{m_i}$ et $s_{i+1}\star$ s'écrit comme la concaténation $u_{m_1} \cdots u_{m_i}$. Par exemple, dans le premier cas : On utilise $|t_i|$ fois la correspondance 2, puis une fois la correspondance 3, puis $|w_i|$ fois la correspondance 2, puis la correspondance 10.

On obtient alors une solution de PCP modifié comme suit :

1	$2^*3\ldots$	$\cdots$						
$\lhd\ q_0\$w\star$	$s_1\star$	$\cdots$	$s_k\star$	$w_k a q_f w'_k\star$	$\cdots$	$w_n q_e\star$	$w_{n-1} q_e\star$	$\cdots$	$\rhd$
$\lhd$	$s_0\star$	$\cdots$	$s_{k-1}\star$	$w_k q_f a w'_k\star$	$\cdots$	$w_n q_f\star$	$w_{n-1} a q_e\star$	$\cdots$	$\$q_e\star\ \rhd$

Réciproquement, si PCP modifié a une solution Alors montrons que M s'arrête sur w.

Soit $u_{i_1} \cdots u_{i_m} = v_{i_1} \cdots v_{i_m}$. On montre, par récurrence sur k qu'il existe un p, et des mots t, t' tels que $u_{i_1} \cdots u_{i_k} = \lhd\ s_1 \star \cdots s_p \star t$, $v_{i_1} \cdots v_{i_k} = \lhd\ s_1 \star \cdots s_{p-1} \star t'$ et ou bien s_p est une configuration finale de la machine, ou bien

— t' est un préfixe de s_p et t est un préfixe de s_{p+1}
— Si t' ne contient pas de symbole d'état et ne se termine pas par $\star$, alors $t' = t$, sinon $t' \vdash_M t$.

Pour $k = 1$, $u_{i_1} = u_1 = \lhd\ q_0\$w\star$ et $v_{i_1} = v_1 = \lhd$. On a bien, pour $p = 1$, $u_{i_1} = \lhd\ s_1\star$, $t = t' = \epsilon$.

Si la propriété est vraie pour k, considérons la paire suivante $u_{i_{k+1}}, v_{i_{k+1}}$. Si s_p était déjà une configuration finale, il n'y a rien à faire. Sinon, $s_p = t' \cdot t'_p$ et $s_{p+1} = t \cdot t_p$. De plus, comme on a une solution de PCP modifié, $v_{i_{k+1}}$ est un préfixe de $t'_p \star t u_{i_{k+1}}$. Considérons successivement tous les $v_{i_{k+1}}$ possibles :

— 1. est impossible car $\lhd$ n'apparait pas dans $t'_p \star t u_{i_{k+1}}$
— 2. Dans ce cas, $t'_p = v_{i_{k+1}} t''_p$ et on a la propriété voulue
— 3, 4,5 : $t'_p = v_{i_{k+1}} t''_p$, t' ne contient pas de symbole d'état, et donc $t = t'$ et $t' \cdot v_{i_{k+1}} \vdash_M t \cdot u_{i_{k+1}}$
— 6,7,8 : comme t'_p ne contient pas $\star$, $t'_p\star = v_{i_{k+1}}$. Donc $s_p\star = t' \cdot v_{i_{k+1}}$. De plus, $t = t'$ et $t \cdot u_{i_{k+1}} = s_{p+1}\star$. Il suffit alors de choisir des mots vides pour les nouveaux t, t'
— 9 : t' doit être vide puisque $v_{i_{k+1}}$ est un préfixe de $t'_p \star t u_{i_{k+1}}$.De plus, on doit avoir $t'_p = \$q_e$, ce qui n'est pas possible. Ce cas n'a pas lieu
— 10. t'_p est vide et il suffit de choisir des mots vides pour les nouveaux t, t'
— 11. Impossible pour des raisons identiques à 9.
— 12,13 : l'état final a été atteint.

Maintenant, reste à montrer que PCP lui-même est indécidable. Pour cela on réduit PCP modifié à PCP comme suit, en supposant (sans perte de généralité) qu'il n'y a pas de paire (ϵ, ϵ).

Si $(u_1, \ldots, u_n)$, $(v_1, \ldots, v_n)$ est une instance de PCP modifié, on considère un alphabet augmenté des lettres $\bullet, \rhd, \lhd$ et l'instance de PCP : $(\lhd \overline{u_1}, \overline{u_1}, \ldots, \overline{u_n}, \bullet \rhd)$, $(\lhd \bullet \widetilde{v_1}, \widetilde{v_1}, \ldots, \widetilde{v_n}, \rhd)$ où $\overline{\epsilon} = \widetilde{\epsilon} = \epsilon$ et $\overline{a \cdot w} = \bullet a \overline{w}$ et $\widetilde{a \cdot w} = a \bullet \widetilde{w}$.

Si PCP modifié a une solution $u_{i_1} \cdots u_{i_m} = v_{i_1} \cdots v_{i_m}$, alors $\lhd \overline{u_{i_1}} \cdots \overline{u_{i_m}} \bullet \rhd = \lhd \bullet \widetilde{v_{i_1}} \cdots \widetilde{v_{i_m}} \rhd$ est une solution de PCP.

Réciproquement, si PCP admet une solution, notons $(u'_0, \ldots, u'_n, u'_{n+1})$ et $(v'_0, \ldots, v'_{n+1})$ l'instance du problème : il existe une suite d'indices telle que $u'_{i_1} \cdots u'_{i_m} = v'_{i_1} \cdots v'_{i_m}$. Notons que, pour tout i, $u'_i = \epsilon$ ou bien u'_i commence par $\bullet$, ou bien $i = 1$. Si $u'_{i_1} = \epsilon$, alors soit k le plus petit indice tel que $u'_{i_k} \neq \epsilon$. Par hypothèse et par construction, $v'_{i_1} \neq \epsilon$ et sa première lettre est $a \notin \{\bullet, \lhd\}$. À l'inverse, la première lettre de u'_{i_k} est dans $\{\bullet, \lhd\}$. Ce qui est absurde. Il en résulte que $u'_{i_1} \neq \epsilon$. Dans ce cas, la première lettre de u'_{i_1} est dans $\{\bullet, \lhd\}$ et donc aussi la première lettre du premier v_{i_k} non vide. Ce n'est possible que si $i_1 = 1$ et cette première lettre est $\lhd$.

Soit maintenant ϕ le morphisme défini sur $(\Sigma \cup \{\bullet, \lhd, \rhd\})^*$ par $\phi(\bullet) = \phi(\lhd) = \phi(\rhd) = \epsilon$ et $\phi(a) = a$ sinon. On montre, par récurrence sur k que $\phi(u'_{i_1} \cdots u'_{i_k}) = u_1 \cdot u_{i_2} \cdots u_{i_k}$ et $\phi(v_{i_1} \cdots v_{i_k}) = v_1 \cdot v_{i_2} \cdots v_{i_k}$, si $1 \leq k < m$. Il en résulte que, si PCP a une solution, alors PCP modifié aussi, en prenant l'image par ϕ.

Exercice 173 (6)
Si $\mathcal{E}$ est un ensemble fini de matrices carrées, le *semi-groupe engendré par* $\mathcal{E}$ est le plus petit ensemble $\mathcal{S}(\mathcal{E})$ qui contient $\mathcal{E}$ et clos par produit : si $M, N \in \mathcal{S}(\mathcal{E})$ alors leur produit MN est dans $\mathcal{S}(\mathcal{E})$.

On veut montrer que le problème suivant est indécidable (*problème de la mortalité de* $\mathcal{E}$) :

Donnée : un ensemble fini de matrices $\mathcal{E}$, à coefficients entiers.

Question : est ce que la matrice nulle est dans le semi-groupe engendré par $\mathcal{E}$?

.

1. Montrer que le problème suivant est indécidable :

   Donnée : un ensemble fini $\mathcal{E}$ de matrices 3×3 à coefficients entiers.

   Question : Existe-t-il une matrice dans $\mathcal{S}(\mathcal{E})$ de la forme $\begin{pmatrix} \alpha & 0 & 0 \\ \beta & 1 & \beta \\ 0 & 0 & \alpha \end{pmatrix}$?

   (**Ind :** On pourra utiliser PCP)
2. Montrer que le problème suivant est indécidable :

   Donnée : un ensembe fini de matrices 3×3,$\mathcal{E}$

   Question : existe-t-il une matrice dans $\mathcal{S}(\mathcal{E})$ dont le coin supérieur gauche est nul ? (i.e. de la forme $\begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$)

3. En déduire que le problème de la mortalité d'un ensemble fini de matrices 3×3 à coefficients entiers est indécidable

Exercice 174 (6)
Montrer que le Problème de correspondance de Post reste indécidable lorsque tous les mots des deux séquences ont pour longueur au plus 2.

Qu'en est il si tous les mots ont pour longueur 2 ?